Competition and critical mass

Maastricht University

Jaap W.B. Bos, Yee Ling Chan, James W. Kolari, Jiang Yuan

Competition and Critical Mass

RM/10/063

METEOR

Maastricht University School of Business and Economics
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

Competition and Critical Mass[☆]

Jaap W.B. Bos[a,*], Yee Ling Chan[b], James W. Kolari[c], Jiang Yuan[b]

[a]*Maastricht University School of Business and Economics, P.O. Box 616, 6200 MD, Maastricht, The Netherlands*
[b]*Utrecht School of Economics, Utrecht University, 3512 BL, Utrecht, The Netherlands*
[c]*Mays Business School, Texas A&M University, TAMU-4218, College Station, Texas 77843-4218, USA*

Abstract

Empirical literature and related legal practice using concentration as a proxy for competition measurement are prone to a fallacy of division, as concentration measures are appropriate for perfect competition and perfect collusion but not intermediate levels of competition. Extending the classic Cournot-type competition model of Cowling and Waterson (1976) and Cowling (1976) used to derive the Hirschman-Herfindahl Index (HHI) of market concentration, we propose an adaptation of this model that allows collusive rents for all, none, or some of the firms in a market. Application of our model and new critical mass measures to data for U.S. commercial banks in the period 1984-2004 confirms that concentration measures are unreliable competition metrics. Our results lead us to conclude that critical mass is a promising new market power metric for competition analyses. Policy and future research implications are briefly discussed.

Keywords: SCP hypothesis, competition, Cournot, conjectural variation, Hirschman Herfindahl index, aggregation bias
JEL: G21, L11, L22

Fallacy of division: The error of assuming that what is true about something must also be true of all or some of its parts.

1. Introduction

Concentration measures, such as the Hirschman-Herfindahl Index (HHI), are commonly used as proxies for competition. Increases in market concentration are believed to increase the potential for collusion, as a negative causal relationship between market concentration and competition is typically assumed (e.g., Cournot-type models). However, both the direction and the validity of the relationship between competition and market concentration can be challenged. A prime example in the industrial organization (IO) literature is the two player Bertrand competitive market model which demonstrates that the number of firms in the market as well as market concentration can be poor proxies for competition (Tirole, 1988).[1] Recent work by Boone (2008) further illustrates the advantages of alternative competition models, and a

[☆]An earlier, now redundant version of this paper appeared as "A Fallacy of Division: The Failure of Market Concentration as a Measure of Competition in U.S. Banking,Ó TKI Working Paper 09-3. We thank Mette Asmild, Jan Boone, Arnound Boot, Keith Cowling, Claire Economidou, Michiel de Jong, Martien Lamers, Ryan van Lamoen, Timothy Riddiough, Stephanie Rosenkranz, Utz Weitzel, Minyan Zhu, participants at the EURO Working Group on Efficiency & Productivity Analysis (EWG-EPA) 2010 International Conference in Chania (Greece), as well as seminar participants at Maastricht University (the Netherlands) and Utrecht School of Economics (the Netherlands) for helpful comments. The usual disclaimer applies.

[*]Corresponding author.
Email addresses: j.bos@maastrichtuniversity.nl (Jaap W.B. Bos), ivy_yeeling@yahoo.com.hk (Yee Ling Chan), j-kolari@tamu.edu (James W. Kolari), j.yuan81@gmail.com (Jiang Yuan)

[1]To some extent the assumptions underlying different IO competition models can be tested. For example, there is extensive literature on the effects of price versus non-price competition on market contestability and entry barriers as well as (tacit) collusion. See Hausman and Sidak (2007), Salop and Scheffman (1983), Kahai et al. (1996), Borenstein (1990), Riordan (1998), Shepherd (1972), Klein (2001), Evans and Kessides (1993), and Martin (1988).

related branch of literature (e.g., Goppelsroeder et al., 2008) documents aggregation problems in constructing market concentration measures. Other studies consider various aspects of concentration measures as proxies for competition by taking into account the notion of strategic behavior among a segment of firms in a market[2] and the existence of large, dominant firms.[3] Particularly relevant to the present paper, this literature highlights the notion of possible dominant firms.[4]

Despite controversy surrounding HHI's usage in competition analyses (White, 2008; Elhauge, 2007), it is routinely applied to mergers by the U.S. Department of Justice. If a merger increases HHI by 100 points, it will likely be under scrutiny by the Antitrust Division.[5] According to horizontal merger guidelines by the Department of Justice and Federal Trade Commission effective since 1992, any market with a post-merger HHI between 1,000 and 1,800 is moderately concentrated, and markets with HHI above 1,800 are concentrated. Recognizing the limitations of market concentration measures, HHI figures are interpreted by authorities in the context of a variety of market conditions, including (for example) the introduction of a new technology through a merger, rise of demand substitutes, growth rate of the market, and entry barriers.

In this paper we demonstrate the limitations of concentration measures as proxies for competition in the context of the classic Cournot-type competition model. By adapting this model for different levels of competition, we obtain *critical mass* as a new measure of market power. Critical mass is defined as the market share at which firms reach market power. A number of related competition measures are manifest, including the percentage of firms with market power, the first year in which at least one firm gained market power, the marginal effect on markup of an increase in market share, the percentage of the markup due to market power, and the dollar value of profits due to rents of colluding firms. Our derivation of critical mass builds upon the seminal work of Stigler (1964) popularized by Cowling and Waterson (1976) and Cowling (1976)(henceforth CW and CL).[6] Extending these studies, we assume that an increase in market share and coincident increase in market concentration *may* result in market power but not necessarily for *all* firms in the market. Hence, although we similarly model and test for Cournot-type competition, we do so allowing for the possibility that collusion is not necessarily either absent or omnipresent. As such, collusive rents may be earned by all, none, or some of the firms in the market. Following Stigler (1964), we assume that collusion becomes more attractive as firms increase their market share due to rising potential losses from not colluding. By modifying the CW and CL model in this way, the critical mass for firms in a market can be derived. Our proposed model incorporates the three classic outcomes of the SCP hypothesis as formulated by CW, CL, and others: (1) for monopoly the critical mass is 100 percent and only one firm has this market share, (2) for perfect competition the critical mass is higher than the highest market share in the market, and (3) for Cournot myopic oligopoly among all firms the critical mass is lower than the lowest market share in the market. Application of our model and new critical mass measure to data for U.S. commercial banks in the period 1984-2004 confirms that concentration measures are unreliable competition metrics. While collusion is prevalent in the banking industry at the state level, the critical mass, rents earned from collusion, and collusive concentration levels vary widely across states. These and other results lead us to conclude that critical mass is a promising new market power metric for competition analyses.

Forthcoming sections develop our proposed model and derivative critical mass measure, present empirical results for the U.S. banking industry, and conclude. An important policy implication of our results

---

[2]See Salop and Scheffman (1983), Evans and Kessides (1993), and Riordan (1998). An excellent review is provided in Scherer and Ross (1990).

[3]See Borenstein (1990), Buschena and Perloff (1991), Kahai et al. (1996), and Klein (2001).

[4]In describing the relation between industry profit rates and concentration, Bain (1951) noted that "... major reliance will have to be placed upon group averaging and upon comparison of group average profit rates at different levels of concentration ..." (p. 309), and "... it is essential in this setting to know whether intragroup variance is of such magnitude as to obliterate the significance of any difference discovered between group averages" (p. 310). As observed by Martin (1988), Bain (1956) posited that small firms would not benefit from market concentration and entry barriers sufficient to support effective collusion, which engenders the possibility of a leading-firm group.

[5]Throughout this paper, we measure HHI as the sum of squared market shares, where the latter are between 0 and a 100, which yields a maximum HHI of 10,000.

[6]For further discussion of HHI, see Adelman (1969), Acar and Sankaran (1999), Kwoka (1998), and others.

is that government merger policies based on concentration measures should be broadened in scope to encompass critical market share and related information.

2. Methodology

CW and CL employed a Cournot model to derive the relation between *industry* markup and HHI. The model's intuitive appeal and straightforward applicability have contributed to its widespread usage. Importantly, these authors assume that the relationship between each firm's markup and market share is the same for all firms in a market, which allows aggregation of these variables. In so doing, empirical estimation of the model is effectively reduced to tests of perfect collusion and perfect competition at the industry level. Only in these two extreme cases is each firm representative of all other firms in the market, and the estimated coefficient for HHI in the test of the resulting SCP relation is either insignificantly different from zero (perfect competition) or unity (perfect collusion). For these two cases there is no fallacy of division. Relevant to the present study, what happens in intermediate cases when the coefficient for HHI lies between zero and unity? Their model implies that there is *some* collusion. But how much? And by whom? Here empirical tests of the model are inconclusive, as the coefficient on HHI (as well as many other concentration measures) can only be compared on an ordinal scale (Kwoka, 1985, 1998; Bikker and Haaf, 2002).

2.1. *The Cournot model revisited*

Revisiting the original Cournot model by CW and CL, each firm chooses its output X_i based on rivals' output levels and seeks to maximize profits Π_i:[7]

$$\begin{aligned} &\Pi_i = pX_i - c_i(X_i) \quad s.t. \\ &p = f(X) = f(\textstyle\sum_{i=1}^{N} X_i), i = 1,2,3,...,N \end{aligned} \tag{1}$$

where $f(X)$ is the inverse demand function. The first order maximization condition of equation (1) is:

$$\frac{\partial \Pi_i}{\partial X_i} = p + X_i f'(X)\frac{\partial X}{\partial X_i} - c_i'(X_i) = 0, i = 1,2,3,...,N \tag{2}$$

where

$$\frac{\partial X}{\partial X_i} = 1 + \frac{\partial \sum_{j=1}^{N} X_j}{\partial X_i} = 1 + \lambda_i, j \neq i \tag{3}$$

and λ_i is the conjectural variation of firm i, which measures the output reaction of its rivals to a change in its own output with $-1 \leq \lambda_{i,t} \leq 1$. Whereas a myopic Cournot oligopoly implies $\lambda_{i,t} = 0$, collusive oligopoly and perfect competition imply $\lambda_{i,t} > 0$ and $\lambda_{i,t} = -1$, respectively. Equation (2) can therefore be rewritten by multiplying the left-hand-side by $\frac{X_i}{X_i}$ and right-hand-side by $\frac{X}{X}$ and then dividing both sides by p to obtain:

$$\frac{pX_i - c_i'(X_i)X_i}{pX_i} = \frac{X_i}{X}\frac{f'(X)X}{p}(1 + \lambda_i), \tag{4}$$

where $c_i'(X_i)X_i$ is the total cost of firm i's output, and $p - c_i'(X_i)X_i$ is firm i's net profit. Thus, the left-hand-side of equation (4) is firm i's markup, also known as the Lerner index L_i (Lerner, 1934). Firm i's market share θ_i is given by $\frac{X_i}{X}$, and the inverse price elasticity of demand $\frac{1}{\eta}$, which is assumed to be the same for all firms, is given by $\frac{f'(X)X}{p}$. Adding time subscripts t, we can now write:

$$L_{i,t} = (-\frac{1}{\eta_t})\theta_{i,t}(1 + \lambda_{i,t}). \tag{5}$$

[7]Without loss of generality, we only consider the total costs of a firm (c_i), rather than its fixed and variable costs.

Thus, each firm's Lerner index depends on the (market) price elasticity of demand, firm market share, and its conjectural variation.

2.2. Proposed model and critical mass

CW and CL make two key assumptions.[8] First, by aggregating equation (5) over all N firms and either assuming that the price elasticity of demand is constant over time or can be captured by control variables, they specify a relationship between industry mark-up and HHI.[9] Subsequent papers have (erroneously) used this result to test for competition by estimating the relationship between firm mark-up and HHI. We will return to this issue in Section 2.4. Second, and relatedly, they crucially assume (as well as many empirical tests of the SCP hypothesis) that conjectural variation $\lambda_{i,t}$ is an omitted variable. In this respect, CW and CL rely upon Stigler (1964), who argued that the (pricing) behavior of firms must be inferred from the way their customers react. From Stigler's rule, we know that "[T]here is no competitive price-cutting if there are no shifts of buyers among sellers ..." (Stigler, 1964, p. 48).

Stigler proved that, given the de facto existence of collusive behavior, the extent to which firms will engage in collusive behavior is directly related to their market share. To see why, following Stigler (1964), assume that a firm targets three groups of customers: new customers, its own old customers, and other firms' old customers.[10] The firm wants to garner its share of the growth of each group.[11] For each group the cost of cheating (i.e., not behaving collusively) is given by the variance of the expected number of customers.[12] The higher this variance, the more likely a firm is to exhibit collusive behavior. Since the three groups are disjoint subsets of the whole customer population, we can simply add up their variances.[13] Consequently, if an increase in market share (θ_i) makes cheating more costly, it will lead to an increase in awareness ($\lambda_{i,t}$) and thereby facilitate collusive behavior. This outcome is supported by the fact that the variance of a firm's expected number of customers increases with an increase in its market share.[14]

Stigler's rule is used by CW, CL, and others to treat a firm's conjectural variation $\lambda_{i,t}$ as an implicit function of its market share. Importantly, the resultant empirical specification only leads to conclusive results if all firms have the same conjectural variation. From equation (5) we can see that this condition holds for two extreme scenarios. First, if all firms behave as myopic Cournot oligopolists, then $\lambda_{i,t} = 0$ for all firms, such that, for a given price elasticity of demand, an increase in θ_i leads to an exactly proportional increase in the Lerner index. Second, in the case of perfect competition, an increase in market share has no impact on performance, as $\lambda_{i,t} = -1$ for all firms. However, returning to our earlier questions, what happens for other intermediate values of $\lambda_{i,t}$? And how much does the impact of an increase in $\lambda_{i,t}$ on the Lerner index depend on a firm's market share?

To answer these questions, we propose the following Lerner index specification:

$$L_{i,t} = \beta_i + \beta_\theta \cdot \theta_{i,t} + \beta_\lambda \cdot \lambda_{i,t} + \beta_{\theta \cdot \lambda} \cdot (\theta_{i,t} \cdot \lambda_{i,t}) + Controls + \varepsilon_{i,t}. \quad (6)$$

[8] See Bikker and Bos (2008).

[9] To be precise, CW include two specifications of their model. In the second specification, they allow for the existence of unequal size firms as determined by their different marginal cost functions. This specification leads to an equality between a slightly modified conjectural variation and HHI. An intuitive way of deriving this result from equation (4) is by multiplying the right-hand-side by $\frac{X_i}{X_i}$, where the denominator of the latter term finds its way into the remainder of the equation, and the numerator (after aggregating the entire equation over N firms) yields HHI instead of firms' market share.

[10] Let: Q_n = number of new customers; Q_o = the total number of old buyers in the market; and q_o^i = the number of old customers for firm i.

[11] First, it wants a share of the new customers (D_n). Second, it wants to retain as many old customers as possible (D_r). And, third, it wants to win over other firms' old customers (D_o).

[12] With the probability of repeat purchases denoted p, the expected number of firm j's customers for each group is given by: $E(D_n^i) = \theta_i * Q_n$; $E(D_r^i) = p * \theta_i * Q_o$; and $E(D_o^i) = (1-p) * \theta_i * (Q_o - q_o^i)$.

[13] A firm expects a consumer to either become a customer (with expectations dependent on its current market share) or not. Thus, for the binomial mean $\mu = n * p$, variance is $n * p(1-p)$, such that variances for each group are given by: $var(D_n^i) = [Q_n * \theta_i * (1-\theta_i)]$; $var(D_r^i) = [Q_o * p * \theta_i * ((1-p)\theta_i)]$; and $var(D_n^i) = [(Q_o - q_0^i) * (((1-p)\theta_i) * (1-(1-p)\theta_i))]$.

[14] In fact: $\frac{\partial var(C_n^i)}{\partial \theta_i} = Q_n - (2 * Q_n * \theta_i) > 0$; $\frac{\partial var(D_r^i)}{\partial \theta_i} = pQ_o - (2 * Q_o * p^2 * \theta_i) > 0$; and $\frac{\partial var(D_o^i)}{\partial \theta_i} = ((1-p)(Q_o - q_0^i)) - (2(1-p) * (Q_o - q_0^i) * \theta_i) > 0$. The first and last equations hold iff $\theta_i < 0.5$. The remaining equation holds iff$p > 2p^2 * M\theta_i$. If $\theta_i < 0.5$, this condition is satisfied also.

Equation (6) differs from previous specifications in two ways. First, we do not aggregate but instead relate the firm's markup to its market share. Second, rather than treating $\lambda_{i,t}$ as an omitted variable, we include it in our empirical specification. In Section 2.3 we elaborate on the measurement and importance of $\lambda_{i,t}$. In Section 2.4, we explain our choice for this particular specification.

For now, to see how equation (6) can be used to reduce the fallacy of division, in line with Stigler's rule, we utilize it to determine how large a firm needs to be to act as a collusive oligopolist. From equation (6) note that:

$$\frac{\partial L_{i,t}}{\partial \lambda_{i,t}} = \beta_{\lambda} + \beta_{\theta\cdot\lambda} \cdot (\theta_{i,t}). \tag{7}$$

Recall that for the oligopolist $\lambda_{i,t} \geq 0$. Therefore, we are interested in knowing at what point $\frac{\partial L_{i,t}}{\partial \lambda_{i,t}} = 0$. Setting the derivative in equation (7) equal to zero and rewriting yields:

$$\theta^{*} = -\frac{\beta_{\lambda}}{\beta_{\theta\cdot\lambda}}, \tag{8}$$

where θ^{*} is the critical mass defined as the market share at or beyond which firms collude. Denoting the lowest (highest) market share in a market by θ^{min} (θ^{max}), we can relate this result to CW and CL by observing that in the case of perfect competition $\theta^{*} > \theta^{max}$, whereas in the case of an oligopoly $\theta^{*} < \theta^{min}$. Moreover, the notion that the likelihood of collusion increases with market share is consistent with Stigler (1964), and $\theta_{i,t} \geq \theta^{*}$ nicely identifies dominant firms (Scherer and Ross, 1990).

Table 1: New competition measures

Abbreviation	Description	Calculation
θ^{*}	*critical mass*	$-\frac{\beta_{\lambda}}{\beta_{\theta\cdot\lambda}}$
$\%^{*}$	*percentage of firms with market power*	$\left(\frac{\sum n \vert \theta_{i,t} \geq \theta^{*}}{\sum n}\right) * 100$
n^{*}	*annual average number of firms with market power*	$\frac{\sum n \vert \theta_{i,t} \geq \theta^{*}}{\text{number of years that year} \geq year^{*}}$
$year^{*}$	*first year in which at least one firm had market power*	$\text{year} \vert \theta_{i,t} \geq \theta^{*}, \text{year}_{-t} \neq year^{*}, \text{for } t = 1, ..., T$
mfx^{*}	*marginal effect of an increase in market share*	$\beta_{\theta} + \beta_{\theta\cdot\lambda} \cdot \lambda_{i,t}$
rents	*percentage of the markup attributable to market power*	$\frac{\left(\frac{\sum \hat{L}_{i,t} \vert \theta_{i,t} \geq \theta^{*}}{\sum n \vert \theta_{i,t} \geq \theta^{*}}\right) - \left(\frac{\sum \hat{L}_{i,t} \vert \theta_{i,t} < \theta^{*}}{\sum n \vert \theta_{i,t} < \theta^{*}}\right)}{\left(\frac{\sum \hat{L}_{i,t} \vert \theta_{i,t} < \theta^{*}}{\sum n \vert \theta_{i,t} < \theta^{*}}\right)} * 100$
dollar bonus	*dollar value of the profits (in thousands) attributable to rents*	$\frac{\text{rents}}{100} * pX_{i,t} - \text{fixed costs}$

For each market the number of firm-year observations included in an estimation is $n = 1, ..., N$, and the number of years is $t = 1, ..., T$.

As shown in Table 1, equation (6) enables new insights into the level of competition in a market. Not only can the critical mass, θ^{*}, be identified, but the percentage of firms which currently operate at or beyond that critical market share, $\%^{*}$, can be estimated. From estimations of equation (6), we can further calculate the number of firms with market power, n^{*}.[15] Also, we can determine the first year in which there was market power, $year^{*}$. Moreover, although the Lucas critique (Lucas, 1976) applies, we may be

[15] This measure is logically similar to the simple concentration ratio obtained by summing the market shares of a subset of the largest firms, C_n, where n is the number of the largest firms.

interested in evaluating the marginal effect of an increase in market share on the Lerner index, mfx^*, to assess the competitive effects of a merger or takeover, or more generally, an increase in relative size. Finally, when there is market power, we can calculate both the rents of colluding firms and the dollar amount of additional profits they derive from their collusive behavior. We do so by assuming that the difference between the average markup of the firms with and without market power is the direct result of collusion.[16]

For policy purposes a more in-depth knowledge of the competitive conditions in a market is often required. When deciding whether or not to approve a merger, the current level of competition is only one piece of required information. The competition measures in Table 1 provide potentially valuable information in this connection. Additionally, such information may be useful to researchers (for example) seeking to explain an inverted-U relationship between competition and innovation (Aghion et al., 2005; Aghion and Griffith, 2005; Aghion et al., 2001). In this case the optimal level of innovation may require a (slightly) positive markup (e.g., see recent U.S. commercial bank evidence by Bos et al., 2009). For industries that are considered to play a key role in the economy, such as the banking sector, there may also be a trade-off between competition and (financial) stability that warrants a (slightly) positive markup (Allen and Gale, 2003). More generally, a variety of other potential applications of the proposed competition measures are readily conceivable.

2.3. *Collusion and critical mass*

In the model of CW and CL, competition is measured at the firm level by each firm's conjectural variation $\lambda_{i,t}$, or the way it expects other firms to react to a change in the size of its operations. Our derivation so far has relied upon the interaction between the firm's current market share $\theta_{i,t}$ and $\lambda_{i,t}$. Hence, given the theoretical setting of this model, we have shown that the assessment of competition brings together both a static market view (as reflected by market shares) and a dynamic market view (as reflected by each firm's conjectural variation). Since these two views are intertwined, as larger firms are expected to perceive different reactions to a change in their output than smaller firms, equation (6) includes the interaction between market share and conjectural variation.

In our interpretation of the marginal effect of an increase in conjectural variation (conditional on market share) on the Lerner index, we relied upon Stigler's (1964) analysis and argued that the likelihood of collusion increases with market share. The latter argument allows us to define critical mass, which represents a dividing line between those that collude and those that do not. In an empirical specification it is desirable to test whether firms with (without) critical mass indeed (do not) collude. As such, we next propose: (1) an estimation strategy, and (2) a way of deriving each firm's conjectural variation $\lambda_{i,t}$ that allows us to distinguish between firms with and without market power.

2.3.1. *Estimation strategy*

In equation (3) we defined each firm's conjectural variation $\lambda_{i,t}$ as $\frac{\partial \sum_{j=1}^{N} X_j}{\partial X_i}, j \neq i$. When estimating equation (6), a correction may be necessary to control for endogeneity, as this equation includes market share measured as $\frac{X_i}{X}$. Without this correction, the estimated coefficients may be biased. In turn, the conditional marginal effect of $\lambda_{i,t}$ on the Lerner index may be inconsistent and therefore unreliable for the purpose of policy making. Also, including the interaction between market share and conjectural variation increases multicollinearity, thereby increasing the standard errors and reducing the likelihood that the estimated coefficient on the interaction term will be significant.

It is obvious that equation (6) may suffer from potential endogeneity due to including both the level of each firm's output (i.e., its market share) and change in output (i.e., its conjectural variation) plus the output levels and changes of all other firms combined. Hence, the treatment of conjectural variation as an

[16]Of course, this approach to calculating both the rents and the dollar bonus is subject to criticism. The most obvious problem is that this comparison does not take into account other ways in which the two groups of firms differ, most notably related to size such as scale economies. However, note that, even in the extreme case in which there are constantly increasing economies of scale, large firms are not guaranteed a higher markup in the case of perfect competition. Rather, large firms would be expected to drive small firms out of the market, in part by undercutting them, thus effectively operating with a *lower* markup.

exogenous regressor is an empirical issue. To address this issue we estimate equation (6) in three alternative ways:

$$
\begin{aligned}
L_{i,t} &= \beta_i + \beta_1 \theta_{i,t} + \beta_2 \lambda_{i,t} + \beta_3(\theta_{i,t}\lambda_{i,t}) + \beta_x(Controls_{i,t}) + \epsilon_{i,t} && \text{(9a)}\\
\Delta L_{i,t} &= \beta_1 \Delta\theta_{i,t} + \beta_2 \Delta\lambda_{i,t} + \beta_3 \Delta(\theta_{i,t}\lambda_{i,t}) + \beta_x \Delta(Controls_{i,t}) + \Delta\epsilon_{i,t} && \text{(9b)}\\
\Delta L_{i,t} &= \beta_1 \Delta\theta_{i,t-1} + \beta_2 \Delta\lambda_{i,t-1} + \beta_3 \Delta(\theta_{i,t-1}\lambda_{i,t-1}) + \beta_x \Delta(Controls_{i,t}) + \Delta\epsilon_{i,t}. && \text{(9c)}
\end{aligned}
$$

The basic specification (9a) is a fixed effect panel estimation that ignores possible endogeneity issues. Specifications (9b) and (9c) instrument for the three key variables in our model: market share, conjectural variation, and the interaction between these two variables. Equation (9b) is estimated in first differences and includes an instrument based on the third and fourth lags of $\theta_{i,t}$, $\lambda_{i,t}$ and $\theta_{i,t}\lambda_{i,t}$. Equation (9c) is also estimated in first differences, and instruments the lags of $\theta_{i,t}$, $\lambda_{i,t}$ and $\theta_{i,t}\lambda_{i,t}$ with the fourth and fifth lags of each variable.

Endogeneity tests for one or more endogenous regressors involve testing the difference between the Sargan-Hansen statistic for the equation with the smaller set of instruments (i.e., equation (9a) in our case) and the equation with the larger set of instruments (i.e., equations (9b) and (9c)). Under the null hypothesis that the specified endogenous regressors are exogenous, this test statistic has a chi-squared distribution.[17] For our purposes, because homoskedasticity is not required, we use the Durbin-Wu-Hausman test.[18]

We begin by estimating both equations (9a) and (9b). If our tests confirm that we cannot reject at least one of our regressors as endogenous, we proceed to test equations (9b) and (9c). Unless otherwise noted, reported results are always based on the first specification that produces unbiased results.

Of course, we are not only concerned about the unbiasedness of the estimated coefficients, but potential multicollinearity associated with the inclusion of the interaction between market share and conjectural variation.[19] According to Brambor et al. (2006, p. 70), multicollinearity raises suspicion when the estimated coefficients in a linear-additive model change due to including an interaction term.[20] However, in our analysis, rather than being interested in the average effect of a variable, we focus on the sign and significance of the conditional marginal effect of conjectural variation on the Lerner index. That is, the significance of the expression in equation (7) is our focal point, not the significance of β_λ and $\beta_{\theta\cdot\lambda}$. From Brambor et al. (2006) we know that the variance of the conditional marginal effect in equation (7) is:

$$
\hat{\sigma}^2_{\frac{\delta L_{i,t}}{\delta\lambda_{i,t}}} = var(\hat{\beta}_\lambda) + \theta^2_{i,t} var(\hat{\beta}_{\theta\cdot\lambda}) + 2\theta_{i,t} cov(\hat{\beta}_\lambda \hat{\beta}_{\theta\cdot\lambda}). \tag{10}
$$

Using equations (7) and (10), marginal effects and their significance can be evaluated at different levels of market share $\theta_{i,t}$. In our empirical analysis, unless otherwise noted, we evaluate the marginal effect at the critical mass $\theta^*_{i,t}$.

2.3.2. Collusion among those that have market power

Do firms with a market share at or above the critical mass behave more collusively than other firms? Following Stigler (1964), we know that collusion is (more) feasible if each colluding firm expects other firms *not* to react to changes in its own output. In other words, among colluding firms, changes in profits (or the Lerner index) result from either changes in its own output or changes in its marginal cost but not from

[17] The degrees of freedom are equal to the number of regressors tested.

[18] As shown by Hayashi (2000, pp. 233-234), the standard test statistic is numerically equal to a Hausman test statistic only under conditional homoskedasticity, whereas the Durbin-Wu-Hausman test is not.

[19] Centering has been suggested as a way of mitigating multicollinearity issues. As pointed out by Brambor et al. (2006) and Kam and Franzese Jr. (2007), centering does not provide us with more accurate data, and "... although the algebraic transformation that results from centering the variables will result in different coefficients and standard errors in the centered model compared to those in the uncentered model, ... this is because they measure different substantive quantities in each model and not because one model produces better estimates than the other" (Brambor et al., 2006, p. 71).

[20] See also Friedrich (1982).

stealing away (losing) earnings from (to) its colluding competitors. Hence, if industry profits change, it is not because profits are reallocated among oligopolists.

We next utilize this view of collusion to once more derive $\lambda_{i,t}$, in order to design an empirical validation test for the critical mass measure $\theta_{i,t}$. We start by considering changes in the industry markup. Letting $\Pi_{i,t}$ ($R_{i,t}$) be profits (total revenues) of firm i at time t, we can write the industry markup at time t as:

$$L_t = \frac{\Pi_t}{R_t} = \frac{\sum_i \Pi_{i,t}}{\sum_i R_{i,t}} = \sum\nolimits_i L_{i,t}\theta_{i,t} = \frac{\sum_i \Pi_{i,t}}{R_{i,t}} \frac{R_{i,t}}{\sum_i R_{i,t}}, \tag{11}$$

where $\Pi_{i,t}$ again denotes firm profits ($pX_i - c_i(X_i)$), and $R_{i,t}$ denotes firm revenues ($pX_{i,t}$). From equation (5) we can write:

$$L_t = \sum\nolimits_i \left[\left(-\frac{1}{\eta_t}\right)\theta_{i,t}(1+\lambda_{i,t})\right]. \tag{12}$$

Also, the change in industry L is:

$$\Delta L_t = \sum\nolimits_i \left[\left(-\frac{1}{\eta_t}\right)\theta_{i,t}(1+\lambda_{i,t})\right] - \sum\nolimits_i \left[\left(-\frac{1}{\eta_{t-1}}\right)\theta_{i,t-1}(1+\lambda_{i,t-1})\right]. \tag{13}$$

At this point we make two additional assumptions. First, we assume that the market price elasticity of demand η is constant. Constant price elasticity is commonly used in empirical demand analyses due to the success of log-linear demand functions (Iwata, 1974, p. 949). Also, empirical studies have found that the market price elasticity of demand is relatively constant (Teles and Zhou, 2005, p. 57).[21] Second, we assume that the conjectural variation of an individual firm is constant in the short run. In the short run firms use historical data to predict their rivals' production and decide on their own output levels. Now we can combine equation (12) with equation (13) to write:[22]

$$\sum\nolimits_i -\frac{1}{\eta}(1+\lambda_i)\Delta\theta_{i,t} = \underbrace{\underbrace{\sum\nolimits_i[\Delta L_{i,t}\cdot\theta_{i,t-1}]}_{\text{operate in t and t-1}}}_{\text{within effect}} + \underbrace{\underbrace{\sum\nolimits_i[\Delta\theta_{i,t}\cdot(L_{i,t-1}-L_{t-1})]}_{\text{operate in t and t-1}} + \underbrace{\sum\nolimits_i[\Delta L_{i,t}\cdot\Delta\theta_{i,t}]}_{\text{operate in t and t-1}} + \underbrace{\sum\nolimits_i[\Delta\theta_{i,t}\cdot L_{t-1}]}_{\text{operate only in t}}}_{\text{reallocation effect}}. \tag{14}$$

Equation (14) clearly captures the effect of collusion on the industry markup. For the myopic Cournot oligolopolist, changes in the industry markup are purely from within each firm. Since firms do not capture market share at the expense of others, changes in the industry markup arise from the within effect (Stiroh, 2000; Stiroh and Strahan, 2003). However, increasing competition leads to increasing reallocation effects, as firms expropriate each others' market shares.[23]

Therefore, in order to derive an expression for $\lambda_{i,t}$, we write equation (14) at the firm level. Removing summations, dividing both sides by $\Delta\theta_{i,t}$, and simplifying further, we can write:[24]

$$-\frac{1}{\eta}(1+\lambda_i) = \Delta L_{i,t}\cdot\frac{\theta_{i,t-1}}{\Delta\theta_{i,t}} + L_{i,t}. \tag{15}$$

where $\frac{\theta_{i,t-1}}{\Delta\theta_{i,t}}$ is the inverse of the percentage change in firm i's market share. Using $\alpha_{i,t}$ to represent the percentage change in firm i's market share at time t, and dividing both sides of equation (15) by $\frac{1}{\eta}$, we can

[21] In fact, our analysis rests on the assumption that the price elasticity of demand is relatively constant over time. Alternatively, the demand function may also be nonlinear. In that case, the marginal benefits from behaving competitively, undercutting competitors and thereby causing a reallocation of profits, vary depending on the shift along the demand curve that results. However, although this would alter the amount of reallocation, it will not affect our analysis otherwise.

[22] A similar type of decomposition is given in Stiroh (2000) and Stiroh and Strahan (2003).

[23] This line of thinking is similar to Caballero and Engel (1993). Note that we use total assets instead of total revenue, as shifts in total revenue capture both changes in rents and changes in output.

[24] The complete derivation is available upon request from the authors.

rearrange to write:[25]

$$\lambda_{i,t} = -\left(\frac{\Delta L_{i,t}}{\alpha_{i,t}} \cdot \eta\right) - (L_{i,t} \cdot \eta) - 1. \tag{16}$$

This expression for $\lambda_{i,t}$ and its lags can be used in estimating equations (9a)-(9c). Also, it can be used to test whether there is more collusion among firms with a market share at or above the critical mass. To do so, we perform three types of tests. First, we test whether reallocation is zero for firms with market power compared to firms without market power. Second, we test whether reallocation equals zero in markets with firms that have market power versus markets that do not have firms with market power. Third, we test whether reallocation is zero for firms with market power before and after they gained market power. Test results are reported in Table 4 and described in Section 3.3.

In sum, we derived a Cournot model that enables identification of firms with market power and construction of related market power measures.

2.4. Conjectural variation as an omitted variable

In Section 2.2, we proposed a Lerner index specification. In line with CW and CL, our specification of the Lerner index in equation (6) includes conjectural variation, $\lambda_{i,t}$, as a covariate for the Lerner index, $L_{i,t}$, in order to capture the direct effects of changes in conjectural variation on firms' markups. And, following Stigler (1964), we need to take into account the role of market share, $\theta_{i,t}$, as a moderator for the effect of conjectural variation on the Lerner Index.[26] It is straightforward to see that the specification in equation (6) satisfies these conditions.

More formally, our specification in equation (6) results from reintroducing, $\lambda_{i,t}$, which has been treated as an omitted variable in most related empirical competition tests, following CW and CL. As a result, existing literature on competition measurement using concentration (or market share) as a proxy suffers from an omitted variable bias. As long as the market structure (or share) variable is biased in the correct way, competition tests reach satisfactory inferences. To clarify this point, Table A.2 in the Appendix summarizes three common specifications. For simplicity, control variables are ignored. In specification *A*, market share is included to proxy for market power, and the dependent variable is the firm-level Lerner index.[27] In common specification *B*, the concentration measure, HHI, is included as a covariate to proxy for market power, and the dependent variable is again the firm-level Lerner index.[28] The last specification *C* is similar to *B*, except that the dependent variable is the industry Lerner index, and the analysis is carried out at the industry-level, as in Cowling (1976) and Cowling and Waterson (1976).

Table A.2 in the Appendix tracks the omitted variable bias for specifications *A*, *B*, and *C*. Starting with the basic specifications in equations (A.1), (B.1), and (C.1), we can introduce $\lambda_{i,t}$, the omitted variable, as a function of the market structure variable using an auxiliary function as shown in equations (A.2), (B.2), and (C.2).[29] In these conditioning equations, $\lambda_{i,t}$ is assumed to have its own firm-specific effect γ_i and depends (positively) on $\theta_{i,t}$.[30] Again introducing $\lambda_{i,t}$ in equations (A.1) and (B.1), and reintroducing λ_t in equation (C.1) results in equations (A.3), (B.3) and (C.3), respectively. We can now assess how accurately the proxies in equations (A.1), (B.1), and (C.1) capture competition by evaluating the omitted variable bias in equations (A.4), (B.4), and (C.4). Ideally, this bias should increase (decrease) with the level of collusion (competition).

[25] As should be expected, rewriting and rearranging equation (5) gives the same result.

[26] "[A] moderator is a qualitative (e.g., sex, race, class) or quantitative (e.g., level of reward) variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable" (Baron and Kenny, 1986, p. 1174).

[27] See Bikker and Bos (2008), and others.

[28] See Berger (1995), Berger et al. (2004), Casu and Girardone (2006), Claessens and Laeven (2004), Gilbert (1984), Molyneux and Forbes (1995), and many others.

[29] In specification *C*, in line with Cowling (1976) and Cowling and Waterson (1976), we use λ_t (the asset-weighted average), rather than $\lambda_{i,t}$.

[30] In specification *C*, consistent with Cowling (1976) and Cowling and Waterson (1976), there is no firm-specific effect, and λ_t is a function of HHI_t.

In specification A, when we estimate a firm-level specification using market share as a proxy for market power, the bias depends on the covariance between $\theta_{i,t}$ and λ_{it} and the variance of $\theta_{i,t}$. From Stigler (1964), we know that the former is positive. The latter reflects the effect of entropy (see Acar and Sankaran, 1999) – as the variance of the market share drops, and firms become more alike, the bias to β_1 (i.e., the coefficient for $\theta_{i,t}$) from using market share as a proxy for competition increases. That is, we can expect to fairly accurately assess the overall level of competition using market share as a proxy, as long as the proxy is highly correlated with conjectural variation, and firms are very alike.

Similarly, with regard to specification C (see Cowling and Waterson, 1976), as long as market concentration (HHI) is highly correlated with conjectural variation (i.e., indicated by the HHI coefficient γ_1 in the conditioning equation), the bias from using market concentration in an industry-level estimation has the correct positive sign. Moreover, the bias increases the more significant this correlation is (as implied by a low variance for γ_1). Although aggregation effects may have (other) negative consequences, the empirical test proposed in Cowling (1976) and Cowling and Waterson (1976) appears to work rather well as long as firms behave alike, viz., when competition is either (close to) perfect, or (almost) absent.

The most interesting specification is B with market concentration included in a firm-level specification, as is common practice in the literature. Equation (B.4) is clearly more complicated than the other specifications. The bias now depends on the covariance between market share ($\theta_{i,t}$) and $\lambda_{i,t}$, the variance of market share, and the market share itself. In fact, equation (B.4) tells us that market concentration works particularly poorly as a proxy for competition in a firm-level specification when concentration is high. To see why, consider the case where we increase the market share by a scaling parameter h.[31] In that case, $Var\left[\theta_{i,t}\right]$ increases by h^2. For a given $Cov\left[\theta_{i,t}, \lambda_{i,t}\right]$, the bias therefore may drop sharply as the increase in market share is sufficiently high. As we shall see in Section 3.4, the result is a high likelihood of both a Type I and II errors when using concentration as a competition measure in this way.

In sum, by treating conjectural variation as an omitted variable and then reintroducing it in equation (6), we can estimate the latter specification in a manner that is both consistent with Stigler (1964) and avoids the fallacy of division. Section 3.5 presents empirical evidence of the bias that exists in U.S. banking competition tests based on specification B.

3. Empirical results

A natural laboratory for examining the relationship between concentration and competition is the U.S. banking industry. Over the last thirty years, historic consolidation has dramatically changed the structure of the banking industry. Studies by Kane (2000), Stiroh and Strahan (2003), Berger et al. (1999), and others document this process, its causes, and some of its potential consequences. There were 15,084 U.S. banking and thrift institutions at year-end 1984 (Jones and Critchfield, 2004, p. 3), but by the end of 2003 the number of institutions had shrank by 48% to 7,842. Numerous studies have sought to determine whether higher bank concentration is detrimental to competition with mixed results. For example, Berger and Hannan (1989) find a positive relationship between profitability and market concentration in retail banking markets in the late 1980s. By contrast, Cole, Goldberg, and White (2004) report no evidence that differences in loan approval procedures of large versus small banks had a negative effect on pricing and volume in the market for small business lending. However, in tests of how competition in local banking markets affects the market structure of nonfinancial sectors, Cetorelli and Strahan (2006) show that potential entrants faced greater difficulty gaining access to credit in concentrated markets than in more competitive markets. Not surprisingly, in most markets characterized by bank consolidation or high market concentration, fears of anti-competitive behavior persist (i.e., more concentrated markets are expected to increase the likelihood of collusion). Berger et al., 2004 provide an excellent survey of the voluminous empirical research on bank market concentration and competition, which was intensely examined in the wake of Depression-era laws

[31] Of course, the sum of the market shares has to equal 1. In practice, the scaling parameter will therefore be negative or zero for some firms. As we are interested in the case where the *average* market share increases over time, we can ignore this additional constraint, as the average effect is measured in equation (B.4).

and regulations that were implemented to recover from massive bank failures in the 1920s and 1930s and promote a smooth functioning financial system.

This section applies our competition model to the U.S. banking industry. After describing the data and estimating our competition indicators, we investigate a number of questions concerning relationships between critical mass and collusion, concentration, rents, and bank deregulation. In general, our results agree with studies that find market concentration constitutes a poor measure of competition (Gilbert and Zaretsky, 2003; Claessens and Laeven, 2004) and, therefore, suffers from a fallacy of division.

3.1. Data

Our data include all insured U.S. commercial banks in the period 1984-2004. We collect year-end Call Report balance sheet and income statement data for individual banks in each state for the period 1984-2004. Empirical tests are conducted on the state level, as most policy decisions involving market concentration measures occur on the state level, evidence supports state-wide pricing (Radecki, 1998; Heitfield, 1999), and studies have found that state-level competition matters more than local competition (Hannan and Prager, 2004; Heitfield and Prager, 2004).[32] As shown by the solid line in Figure 1, state level market concentration measured by HHI rose considerably during the period under consideration. At the same time the dashed line shows that the number of banks was almost halved.

Figure 1: Consolidation in U.S. banking

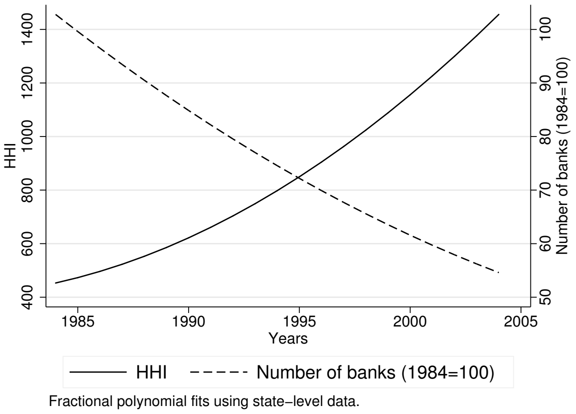


Table 2 contains the descriptive statistics for our variables. Markup is calculated as the sum of profit before tax and fixed asset expenditures over total revenues. Market share is based on total assets, and conjectural variation is calculated via equation (16) using the markup, market share, and market price elasticity of demand. Following empirical methods in Teles and Zhou (2005), we estimated the latter elasticity to be -0.15.[33] In the sample period 1984-2004, the average markup $L_{i,t}$ was close to 16%, the average market share was 0.5%, and average conjectural variation was -0.943.

When estimating equations (9a)-(9c), consistent with previous literature (Bikker and Bos, 2008), we utilize control variables. To take into account differences in risk-taking, we include the ratio of total loans and leases over total assets as a proxy for credit risk exposure. We also include earning assets measured as the

[32] Of course, banks can compete on local, state, and national levels to various degrees (e.g., see Gilbert and Zaretsky, 2003).

[33] Teles and Zhou (2005) proxy the market price elasticity of demand with the interest elasticity of money demand. The real money demand function is estimated for our sample period. The dependent variable is the deflated MZM money aggregate (i.e., M2 less small time deposits plus institutional money market funds), and the independent variables are GDP and interest rate (i.e., the federal funds rate minus the MZM own rate). The first derivative of this specification and one-period lagged dependent variable is used to adjust for serial correlation. Based on this model, we estimate the interest elasticity of money to be -0.15, which is close to the -0.20 estimate obtained by Teles and Zhou (2005, p. 57) in the similar sample period 1980-2003.

ratio of total assets minus fixed assets over total assets.[34] The ratio of noninterest expenses over noninterest income is used to control for costs, rather than the ratio of total expenses over total income, as year-to-year changes in interest earnings and expenses may reflect yield curve changes instead of banks' cost management.[35] Finally, to control for the fact that market shares as well as markups may reflect differences in efficiency (i.e., the so-called efficiency hypothesis (Goldberg and Rai, 1996)), we include cost efficiency estimates based on standard translog cost frontier methods (see Bos and Kolari (2005) and citations therein).

Table 2: Descriptive statistics for variables

Variable		Mean	Std. Dev.
markup	($L_{i,t}$)	0.159	0.225
market share	($\theta_{i,t}$)	0.005	0.026
conjectural variation	($\lambda_{i,t}$)	1.000	0.021
risk	(loans and leases/total assets)	0.557	0.152
earning assets	(total assets - fixed assets)/total assets)	0.983	0.013
cost	(noninterest expenses/noninterest income)	5.778	9.795
cost efficiency	(translog cost frontier estimates)	0.791	0.096

The total number of observations based on the specification in equation (9a) is 200,488. The p-value for the one-sided t-test that $\lambda_{i,t} > 1$ (<) equals 0.828 (1.000). Annual cost efficiency estimates were obtained from Bos and Kolari (2005), who employed a translog cost function to obtain estimates.

With the exception of the cost ratio, the standard deviations of the control variables are relatively low. This low variability may (in part) explain why later empirical results suggest that control variables play a minor role. In order to test whether our assumption holds that differences in markup are not attributed to differences in the control variables, we identify banks with and without market power and run a canonical linear discriminant analysis using all four control variables. Results show that the probability of correctly classifying a bank with (without) market power based on the control variables is 60.06% (56.10%), confirming that the control variables indeed are not driving our results.

Another reason for this result is that our model is grounded in an identity and controls for firm-level heterogeneity either through fixed effects using equation (9a) or dynamic panel estimators using equations (9b) and (9c).

3.2. *Main empirical results*

Table 3 summarizes our competition measures for states in which there is evidence of collusion. Endogeneity appears to be less of an issue than was originally believed. Although we find evidence of collusion in 30 states, tests indicate that instrumenting is required in only 12 states, and estimating specification (9c) was never warranted. For illustrative purposes, Figure 2 provides a graphical representation of the estimation results for California. The graph shows the marginal effect of a change in conjectural variation on the markup conditional on market share. Since our model posits that collusion results in a marginal effect equal or greater than zero, we can infer from the results in Figure 2 that collusion exists in California. As shown in Table 3, we estimate California's critical mass $\theta^*_{i,t}$ to gain market power at slightly more than 6 percent. Other results in the table indicate that on average less than 1 percent of California banks had market power. Note that collusion first occurs in 1989 when state-level HHI was a modest 1,074. Also, approximately three banks had market power and earned rents averaging almost 58 percent of the total markup.

Our results for California aptly demonstrate the fallacy of division. Although the market is moderately concentrated, there is collusion among a very small number of banks. In effect, according to our results, this

[34] Additionally, we tested the leverage ratio equal to total equity over total assets with little change in results.
[35] In unreported results, the inclusion of total expenses over total revenues did not alter our conclusions.

Table 3: Empirical results for states with market power

State	*Critical mass* θ^*	*Banks with market power (%)* $\%^*$	*First year with market power* $year^*$	*HHI in first year with market power* HHI^*	*Marginal effect of change in market share* mfx^*	*Markup due to market power (%)* rents	*Profits due to market power* bonus	*Average number of banks with* θ^* n^*	*Instrument test (p-value)* endog.	*Specification* spec.
Alabama	1.73***	2.60	1984	571.79	2.49	24.97	227141.30	5.13		eq (9a)
Alaska	44.21***	11.01	1989	3325.73	1.61	13.74	14936.30	0.44		eq (9a)
California	6.06***	0.72	1989	1074.26	1.35	57.80	3088269.00	3.06		eq (9a)
Colorado	0.47***	15.15	1984	234.77	8.42	9.34	1225.39	43.98		eq (9a)
Dist. of Columbia	0.00***	100.00	1984	2403.24				17.76		eq (9a)
Florida	0.22***	21.39	1984	364.69	13.93	84.29	76863.87	69.44	0.04**	eq. (9b)
Georgia	5.88***	0.66	1984	723.39	18.51	94.55	1354554.00	2.45	0.00***	eq. (9b)
Hawaii	0.00***	100.00	1984	2767.40				17.30		eq (9a)
Idaho	8.11***	41.27	2001	895.53	2.36	15.46	13126.19	4.00	0.06*	eq. (9b)
Illinois	1.99***	0.83	1998	888.39	1.97	20.94	304911.60	5.41		eq (9a)
Indiana	1.19***	5.49	1984	209.35	0.27	21.04	42393.09	13.36		eq (9a)
Iowa	0.18***	28.64	1984	51.57	0.50	9.33	683.41	144.67		eq (9a)
Kansas	13.44***	0.09	1991	259.36	84.92	102.75	406087.40	0.24	0.00***	eq. (9b)
Missouri	0.46***	8.98	2000	395.28	0.57	16.18	12531.79	27.48		eq (9a)
Montana	1.82***	8.40	1984	149.95	0.97	32.13	9531.43	9.93		eq (9a)
New Mexico	18.23***	0.91	1985	466.91	39.92	69.63	153705.10	0.61	0.02**	eq. (9b)
New York	0.02***	59.41	1990	1021.53	0.01	7.83	18723.18	96.02		eq (9a)
Oklahoma	1.00***	4.12	1994	300.15	41.05	97.61	97043.93	10.38	0.00***	eq. (9b)
Oregon	0.00***	100.00	1984	2751.76				47.67	0.03**	eq. (9b)
Pennsylvania	0.00***	100.00	1984	620.41				254.98	0.00***	eq. (9b)
South Carolina	5.96***	5.22	1995	1372.66	1.46	30.63	82510.63	3.48		eq (9a)
South Dakota	19.05***	1.88	2003	2019.01	5.25	76.75	994524.40	1.10	0.00***	eq. (9b)
Tennessee	1.07***	3.07	1994	627.30	1.10	19.58	62170.29	9.63		eq (9a)
Utah	21.05***	3.31	1984	1589.84	1.45	23.06	158612.10	1.62		eq (9a)
Virginia	0.06***	88.42	1984	1162.51	0.15	28.03	7560.33	138.11		eq (9a)
Washington	0.58***	27.27	1984	1710.37	0.38	19.19	10362.43	22.58		eq (9a)
West Virginia	0.00***	100.00	1984	101.41				160.01	0.01***	eq. (9b)
Wisconsin	2.89***	0.81	1986	118.88	43.33	101.87	458235.20	2.96	0.00***	eq. (9b)

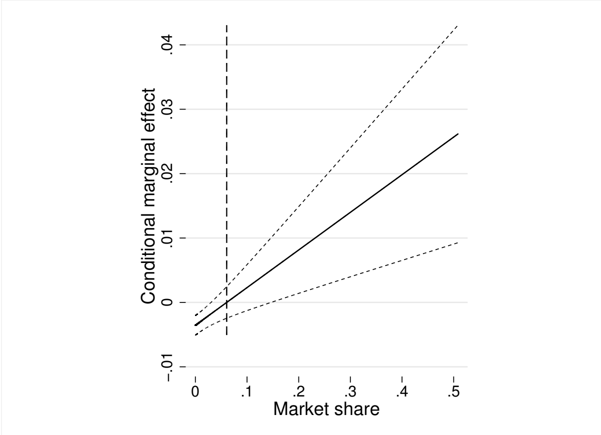
Results are based on the preferred specification denoted in the last column. Significance is indicated at the following levels: 1/5/10% (***/**/*, respectively).

market's relevant concentration measure is a C_3 ratio. Even though rents for these top 3 banks are sizeable, our results suggest that the label *collusive* does not fit most banks operating in the California market.

The estimated competition metrics in Table 3 reveal that, among the states in which there is evidence of collusion, the HHI at which collusion begins differs widely. In Arkansas, Iowa, Montana, Oklahoma, and West Virginia, this minimum HHI is always less than 150. By contrast, in Alaska, the District of Columbia, Hawaii, and Oregon, the minimum HHI is always above 2,500. Consequently, HHI as a measure of competition appears to be seriously flawed. This problem is further illustrated by the critical mass required to have market power, which ranges from 44.21 percent in Alaska to 0 percent in Pennsylvania. The use of other market concentration measures, such as a C_5 or C_{10} ratio, appears to be subject to question also, as the average number of banks with market power ranges (for example) from less than 1 in South Dakota to more than 269 in Kentucky. Likewise, rents widely range from a modest 7.49 percent in Mississippi to slightly more than 100 percent in Arkansas, Kansas, and Wisconsin.[36] The first year in which collusion

[36]Results for New Mexico are suspect due to estimated rents of -252.03 percent, or a negative dollar bonus.

Figure 2: Conditional marginal effect of conjectural variation on the markup in California



started has a fair range too. In 20 out of 30 states, collusion already existed in the 1980s, often from the beginning of our sample period. The last year in which collusion started is 2004, when banks in New Mexico reached their critical mass.

In an effort to verify these results, we estimated specifications (9a)-(9c) for four separate periods: 1984-1988, 1989-1993, 1994-1998 and 1999-2004. Table A.1 in the Appendix reports the results for states and periods in which there is evidence of collusion. For the states included in Table 3, the results are generally robust when compared to Table A.1. However, the latter table also includes a number of other states in which there is evidence of collusion in one or more of the four periods. In the remainder of this section, we focus on the main results in Table 3, except when the distinction between periods is crucial, as in the case of the Interstate Banking and Branching Efficiency Act of 1994.

In sum, although there is substantial evidence of collusion at the state level in U.S. banking, the extent to which banks colluded, the rents they earned from colluding and, importantly, the level of concentration at which collusion started for (a subset of) banks varied widely across states. These results confirm the dangers of using market concentration as a singular measure of competition.

3.3. Does collusion increase above the critical mass?

In section 2.3.2 we distinguished between banks that operate with and without market power and proposed three related hypotheses. Table 4 contains the test results for these hypotheses.[37]

The first and most direct market power test compares banks with market power to banks without market power. Our results confirm that the reallocation effect in equation (14) equals zero for banks with market power but not for banks without market power. The second test of whether the reallocation effect equals zero in states where banks have market power versus states where no banks have market power yields similar results. The third, and last, test of whether the reallocation effect equals zero in states where banks have market power before and after the first year in which at least one bank gained market power again yields similar results. Together, these test results suggest that collusion does increase above the critical market share. In this regard, collusion can take place among a select number of firms in a market. Also, banks that are too small and do not collude experience a markup change due to reallocation effects (as banks gain market share at the expense of other firms with higher markups).[38]

[37]In addition to the tests reported in Table 4, we performed unreported tests. First, non-parametric rank tests yielded qualitatively identical results. Second, we scaled the reallocation effects by the size of the banks, which also did not change the results. In our opinion the reported results constitute a strong test of our hypothesis, as scaling especially reduces the reallocation effects of large banks.

[38]The implication here is a negative reallocation effect.

Table 4: More collusion results in less reallocation

Tests	Null hypothesis	p-value	Result
Banks with market power versus	$reallocation_{(\theta_{i,t} \geq \theta^*_{i,t})} = 0$	0.506	accept
banks without market power	$reallocation_{(\theta_{i,t} \geq \theta^*_{i,t}} = 0$	0.002	reject
States with market power versus	$reallocation_{(year \geq year^*)} = 0$	0.436	accept
states without market power	$reallocation_{(year < year^*)} = 0$	0.000	reject
Banks with market power before year versus*	$reallocation_{(year \geq year^*)\|year^* \neq .} = 0$	0.334	accept
*after year**	$reallocation_{(year < year^*)\|year^* \neq .} = 0$	0.017	reject

All p-values are based on two-sided t-tests with a critical value of 5%.

3.4. *Additional empirical results*

In this section we further investigate relationships between concentration measures and market power (competition) metrics by means of graphical and empirical analyses.

3.4.1. *Is there more collusion when the market is more concentrated?*

Implicit in HHI as a measure of competition is the idea that there is a higher probability of collusion when a market is more concentrated. Our model makes clear that this assertion may be flawed. Although the probability of collusion increases in our model with higher HHI, it does not necessarily imply that the number (or percentage) of banks colluding increases. Additionally, because a higher HHI often implies fewer banks, endogeneity may be an issue. As the results in Table 4 show, this does not necessarily mean that all remaining banks in a market collude. By way of illustration, Figure 3a compares the percentage of banks with market power ($\%^*$) with the HHI in the first year in which there was market power (HHI^*). To simplify the graphical analysis, the sample is restricted to states in which $\%^*$ was at least 5 percent.

If the intuition for HHI as a measure of competition is correct, we expect to see a higher percentage of banks with market power in states with higher HHIs. However, the results shown in Figure 3a do not support this relation, as casual inspection suggests a slightly negative association between market power and HHI.

Certainly the number of observations in Figure 3a is low, and the results may be affected by the fact that we only consider the first year in which there is collusion. For these reasons a more formal analysis is provided in the first part of Table 5. Using the results in Table A.1, we regress the percentage of banks with market power on HHI. Again we find no positive relationship between these variables, which contradicts the intuition underlying HHI as a competition measure. Consequently, an important potential pitfall is Type I error in terms of ignoring collusion in markets with low concentration.

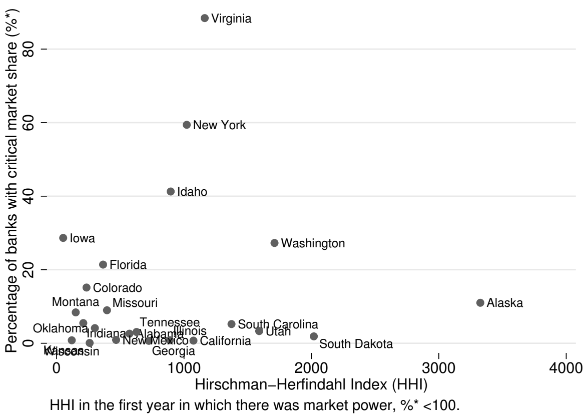
3.4.2. *Is the critical mass lower in more concentrated markets?*

As noted in the introduction, mergers are more likely to be approved in markets with lower concentration. Our results so far raise serious questions about this practice, as many states have low concentration levels but exhibit low critical mass. As such, there is the possibility of a Type II error by authorities in terms of approving a merger even though it leads to (or reaffirms) market power.

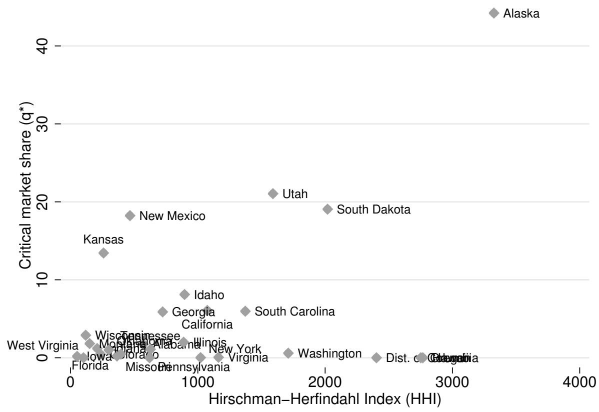
An important caveat is that, consistent with the Lucas critique, a merger may alter the market dynamics. To partially address this concern, consider the relationship in Figure 3b between critical mass and HHI in the first year in which there is market power. This relationship appears to be positive. For the states shown there, based on casual observation of the marginal effects of increases in market share on markup, an increase in market share (e.g., due to a merger) tended to have a much larger effect in less concentrated markets. This relation is confirmed by the regression results in Table 5, in which a positive estimated

Figure 3: The fallacy of division

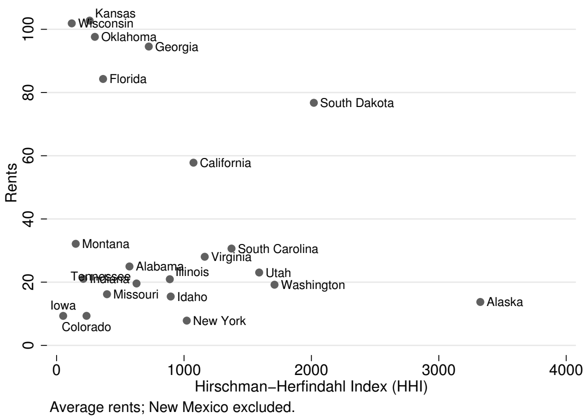
(a) Is there more collusion in more concentrated markets?



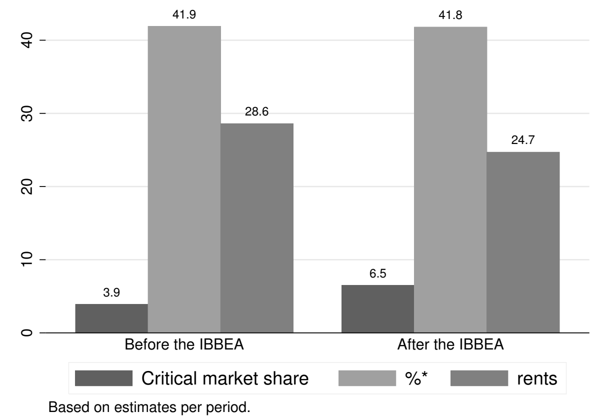
(b) Is the critical mass lower in more concentrated markets?



(c) Are there (more) rents for banks in more concentrated markets?



(d) Has competition increased after the IBBEA of 1994?



coefficient is found when critical market shares are regressed on HHI. Once again, there is a risk of Type II errors, as authorities may be inclined to wrongly approve mergers in states with low concentration, even though the merger may result in significant market power.

3.4.3. Do banks in more concentrated markets earn higher rents?

The extent to which regulatory authorities contemplate Type I and II errors in assessing the level of competition depends on how market power affects consumers. As noted earlier, some market power may foster financial innovation and stability. Therefore, it is important to know the extent to which market power enables firms to capture rents. Figure 3c shows the relationship between our rough measure of rents earned by banks with market power and HHI. The overall pattern of the data is disconcerting, as banks with high rents appear to operate in markets with low concentration.[39] However, results from regressing rents on HHI in Table 5 do not suggest a significant relationship between these two variables. We infer that market concentration likely has little effect in terms of enabling those with market power to earn rents.

[39] Based on Table 4, these results may reflect the low number of observations for banks with market power to some degree.

Table 5: Regressions of market power measures on HHI in states

Is there more collusion when the market is more concentrated?					
	HHI		Constant		*N*
%*	0.004	(0.83)	43.853	(6.17)**	95
Is the critical mass lower in more concentrated markets?					
	HHI		Constant		*N*
$\theta^*_{i,t}$	0.002	(2.54)*	-0.268	(0.19)	95
Do banks in more concentrated markets earn higher rents?					
	HHI		Constant		*N*
rents	0.014	(1.53)	9.738	(0.62)	67

Estimations for the first and second question are based on a Tobit regression model with upper (100) and lower (0) limits. Estimations for the third question are based on a standard regression model. All estimations utilize estimated results from Table A.1.

3.4.4. Has competition increased after the IBBEA of 1994?

In order to stimulate competition, legal and regulatory changes aimed at lowering entry barriers, removing geographic barriers, lowering costs, and enabling innovation have been periodically implemented in the U.S. banking industry. A major legislative change in our sample period took place in 1994 when the Interstate Banking and Branching Efficiency Act (IBBEA) was passed by the U.S. Congress. This Act sought to significantly enhance competition by allowing banks to compete across state borders. In view of major restructuring in the banking industry due to this legislation, as a final assessment of the level of competition in U.S. banking, we briefly examine the impact of the IBBEA.

Table 6: Competition before and after the 1994 IBBEA

States with collusion before	37
States with collusion after	36
States with collusion before and after	26
States with collusion only *before*	11
States with collusion only *after*	10

Based on estimates per period provided in Table A.1.

Dividing our sample data into the subperiods 1984-1993 and 1994-2004, Figure 3d summarizes different aspects of collusion before and after the IBBEA. The critical mass required to have market power increased from an average of about 4 percent to about 6.5 percent. However, due to increases in concentration, the percentage of banks with market power (in states where there is collusion) did not change. Also, average rents earned decreased by almost 4 percentage points. Table 6 provides further information (based on Table A.1.). The number of states with collusion stayed almost constant before and after IBBEA. Although there were 11 states with collusion only before interstate banking deregulation, 10 states experienced collusion only after the IBBEA was implemented. Thus, we infer that, while the IBBEA had mixed effects on bank competition across states, on average competition was not generally affected on the state level.

3.5. How important is omitted variable bias?

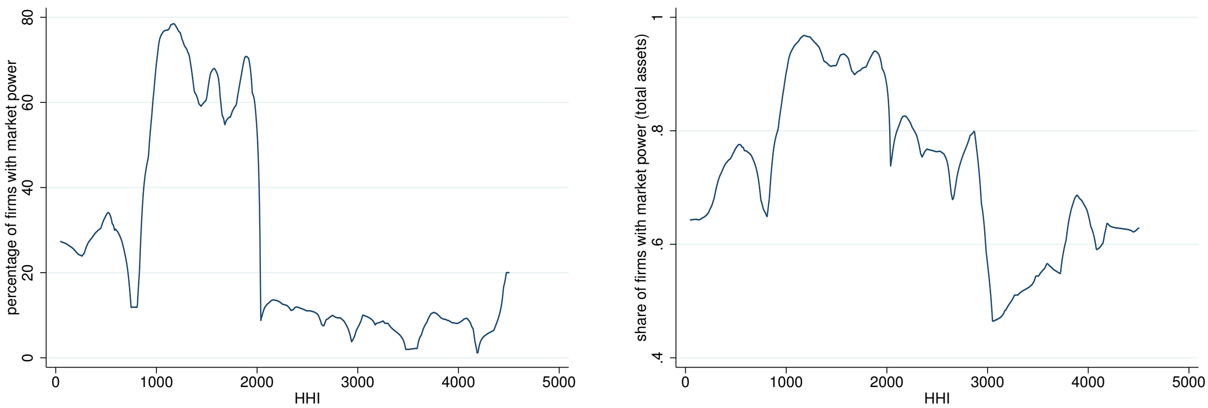
In Section 2.4 and Table A.2, we analyzed the bias that exists when conjectural variation is omitted from Cournot-type competition tests. The theoretical evidence presented there demonstrated that the most serious bias exists when we relate firm-level performance (i.e., the Lerner index) to market concentration (i.e., the HHI). In the latter case, a combination of omitted variable bias and aggregation bias means that HHI is a particularly poor proxy for the level of collusion when it is sufficiently high. But what is *sufficient* in the context of our empirical analysis?

We next provide an empirical answer to this question by relating HHI to the percentage of firms with market power, %*. As a further robustness test, we relate HHI to the percentage of the bank assets held by banks with market power. For each of these measures, we run a kernel regression, where HHI is the explanatory variable. The intuition is that, as long as collusion increases with HHI, the conditional densities for our measures that result from the kernel regression should follow a positive relationship. Once this positive relationship no longer holds, as described in Section 2.4, market concentration becomes a poor proxy for competition in a firm-level specification.

Figure 4: Omitted variable bias and the HHI

(a) Percentage of firms with market power

(b) Percentage of assets with market power



Figures 4a and 4b show the predicted percentages of firms (assets) with market power as HHI increases. In line with equation (B.4) in Table A.2, market concentration works particularly poorly as a proxy for competition in a firm-level specification when concentration is high. Initially, as HHI increases so does our market power measure, from an HHI of 0 to about 2000. Then an inverse relation between HHI and market power is observed as HHI exceeds 2000, even though at higher HHIs we are supposed to observe more collusion. Paradoxically, HHI is weakest when we need to rely on it the most for information about market power.

4. Conclusion

Competition tests based on market concentration measures, such as the popular Hirschman-Herfindahl Index (HHI), assume either perfect competition or perfect collusion and, therefore, are prone to a fallacy of division. In the real world there is a continuum between these extreme endpoints that encompasses many competitive market conditions. Extending previous work by Cowling and Waterson (1976) and Cowling (1976), we relax this restrictive assumption by assuming that Cournot-type collusive rents can be earned by all, none, or some of the firms in a market. We propose that a firm's markup (or Lerner index) is a function of its market share, conjectural variation (or reaction of rivals to a change in a firm's output), and their interaction. The proposed model allows estimation of how large a firm must be to achieve *critical mass* as a collusive oligopolist. Using critical mass, the percentage of firms with market power can be estimated. Other potentially valuable market power metrics are the first year in which at least one firm gained market power, the marginal effect on markup of an increase in market share, percentage of the markup due to market power, and the dollar value of profits due to rents of colluding firms.

Empirical specifications of the model are developed that take into account the possibility of endogeneity as well as a new measure of conjectural variation. We applied the model to data for U.S. commercial banks at the state level in the period 1984-2004. In general, our results do not support concentration as a singular measure of competition. While we found considerable evidence of collusion at the state level, the extent of collusion, rents earned from collusion, and collusive concentration levels varied widely across states. Even when some banks had market power in a state, it was rarely the case that collusion was perfect among all banks in the state. Instead, large banks in some states colluded. Moreover, reliance on concentration as a measure of competition results in both Type I and II errors, as there were both states with high-concentration/high-competition levels and states with low-concentration/low-competition levels. Consistent with our model, we found more collusion among banks with market shares beyond the critical mass required in a state. However, there was no clear time pattern as to when a state reaches the critical mass. Compared to less concentrated markets, more concentrated markets did not exhibit more collusion, tended to have larger critical mass, and were not more likely to earn collusive rents. Lastly, interstate banking deregulation did not have major competitive effects on the banking industry.

A number of previous studies address weaknesses in concentration as a competition proxy. We contribute to this literature by proposing a Cournot model solution that traces its roots back to the original, and frequently misinterpreted, work of Cowling and Waterson (1976) and Cowling (1976). In addition, we suggest a way of validating our proposed alternative critical mass competition measures. Lastly, we demonstrate, both theoretically and empirically, that HHI is a particularly poor proxy for competition when we need to rely on it the most.

We conclude that critical mass is a promising new market power metric for competition analyses. It constitutes a more robust measure of market power than HHI that incorporates added information about the degree of competition. As such, it is more general than concentration measures, which are are only accurate in the extreme cases of perfect competition and perfect collusion, rather than more realistic intermediate levels of competition. An important policy implication is that the U.S. Department of Justice, regulatory agencies, and similar authorities in other countries should supplement HHI concentration guidelines with information on critical mass and related market power metrics. In this way a more complete picture of market competition and collusion can be obtained in merger analyses and decisions. Finally, our new competition measures open up a number of new avenues for possible future study. For example, further research is recommended with respect to critical mass analyses of other industries, definition of the relevant market, validity of Cournot-type competition models, and other economics applications of critical mass.

References

Acar, W., Sankaran, K., 1999. The myth of the unique decomposability: Specializing the Herfindahl and entropy measures? Strategic Management Journal 20 (10), 969–975.

Adelman, M. A., 1969. Comment on the "H" concentration measure as a numbers-equivalent. The Review of Economics and Statistics 51 (1), 99–101.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: An inverted-U relationship. Quarterly Journal of Economics 120 (2), 701–728.

Aghion, P., Griffith, R., 2005. Competition and Growth: Reconciling Theory and Evidence. MIT Press, Cambridge, MA.

Aghion, P., Harris, C., Howitt, P., Vickers, J., 2001. Competition, imitation and growth with step-by-step innovation. Review of Economic Studies 68 (3), 467–492.

Allen, F., Gale, D., 2003. Comparing Financial Systems. MIT Press, Cambridge, MA.

Bain, J. S., 1951. Relation of profit rate to industry concentration: American manufacturing, 1936-1940. The Quarterly Journal of Economics 65 (3), 293–324.

Bain, J. S., 1956. Barriers to New Competition. Harvard University Press, Cambridge, Massachusetts.

Baron, R. M., Kenny, D. A., 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51 (6), 1173–1182.

Berger, A. N., 1995. The profit-structure relationship in banking–tests of market-power and efficient-structure hypotheses. Journal of Money, Credit and Banking 27 (2), 404–431.

Berger, A. N., Demirguc-Kunt, A., Levine, R., Haubrich, J. G., 2004. Bank concentration and competition: An evolution in the making. Journal of Money, Credit and Banking 36 (3), 433–451.

Berger, A. N., Demsetz, R. S., Strahan, P. E., 1999. The consolidation of the financial services industry: Causes, consequences, and implications for the future. Journal of Banking and Finance 23 (2-4), 135–194.

Berger, A. N., Hannan, T. H., 1989. The price-concentration relationship in banking. The Review of Economics and Statistics 71 (2), 291–299.

Bikker, J. A., Bos, J. W. B., 2008. Bank Performance: A Theoretical and Empirical Framework for the Analysis of Profitability, Competition and Efficiency. International Studies in Money and Banking. Routledge, London and New York.
Bikker, J. A., Haaf, K., 2002. Competition, concentration and their relationship: An empirical analysis of the banking industry. Journal of Banking and Finance 26 (11), 2191–2214.
Boone, J., 2008. A new way to measure competition. The Economic Journal 118 (531), 1245–1261.
Borenstein, S., 1990. Airline mergers, airport dominance, and market power. The American Economic Review 80 (2), 400–404.
Bos, J. W. B., Kolari, J. W., 2005. Large bank efficiency in Europe and the United States: Are there economic motivations for geographic expansion in financial services? Journal of Business 78 (4), 1555–1592.
Bos, J. W. B., Kolari, J. W., van Lamoen, R. C. R., 2009. Competition and innovation: Evidence from financial services. TKI Working Paper 09-16, Tjalling Koopmans Institute, Utrecht School of Economics.
Brambor, T., Clark, W., Golder, M., 2006. Understanding interaction models: Improving empirical analyses. Political Analysis 14(1), 63–82.
Buschena, D. E., Perloff, J. M., 1991. The creation of dominant firm market power in the coconut oil export market. American Journal of Agricultural Economics 73 (4), 1000–1008.
Caballero, R. J., Engel, E. M. R. A., 1993. Microeconomic rigidities and aggregate price dynamics. European Economic Review 37 (4), 697–711.
Casu, B., Girardone, C., 2006. Bank competition, concentration and efficiency in the single European markets. The Manchester School 74 (4), 441–468.
Cetorelli, N., Strahan, P. E., 2006. Finance as a barrier to entry: Bank competition and industry structure in local U.S. markets. Journal of Finance 61 (1), 437–461.
Claessens, S., Laeven, L., 2004. What drives bank competition? some international evidence. Journal of Money, Credit and Banking 36 (3), 563–583.
Cole, R. A., Goldberg, L. G., White, L. J., 2004. Cookie cutter vs. character: The micro structure of small business lending by large and small banks. Journal of Financial and Quantitative Analysis 39 (02), 227–251.
Cowling, K., 1976. On the theoretical specification of industrial structure-performance relationships. European Economic Review 8 (1), 1–14.
Cowling, K., Waterson, M., 1976. Price-cost margins and market structure. Economica 43 (171), 267–274.
Elhauge, E. R., 2007. Harvard, not Chicago: Which antitrust school drives recent supreme court decisions? Harvard Law and Economics Discussion Paper No. 594.
Evans, W. N., Kessides, I. N., 1993. Localized market power in the U.S. airline industry. The Review of Economics and Statistics 75 (1), 66–75.
Friedrich, R. J., 1982. In defense of multiplicative terms in multiple regression equations. American Journal of Political Science 26 (4), 797–833.
Gilbert, R. A., November 1984. Bank market structure and competition: A survey. Journal of Money, Credit and Banking 16 (4), 617–44.
Gilbert, R. A., Zaretsky, A. M., 2003. Banking antitrust: Are the assumptions still valid? Federal Reserve Bank of St. Louis Review 85 (6), 29–52.
Goldberg, L. G., Rai, A., 1996. The structure-performance relationship for European banking. Journal of Banking & Finance 20 (4), 745–771.
Goppelsroeder, M., Schinkel, M. P., Tuinstra, J., 2008. Quantifying the scope for efficiency defense in merger control: the Werden-Froeb-Index. The Journal of Industrial Economics 56 (4), 778–808.
Hannan, T. H., Prager, R. A., 2004. The competitive implications of multimarket bank branching. Journal of Banking and Finance 28 (8), 1889–1914.
Hausman, J. A., Sidak, J. G., 2007. Evaluating market power using competitive benchmark prices rather than the Hirschman-Herfindahl index. Antitrust Law Journal 74 (2), 387–408.
Hayashi, F., 2000. Econometrics, 1st Edition. Princeton University Press., Princeton, NJ.
Heitfield, E., Prager, R. A., 2004. The geographic scope of retail deposit markets. Journal of Financial Services Research 25 (1), 37–55.
Heitfield, E. A., 1999. What do interest rate data say about the geography of retail banking markets? Antitrust Bulletin 44, 333–347.
Iwata, G., September 1974. Measurement of conjectural variations in oligopoly. Econometrica 42 (5), 947–966.
Jones, K. D., Critchfield, T., 2004. The declining number of U.S. banking organizations: Will the trend continue? Federal Deposit Insurance Corporation Future of Banking Study FOB-2004-02.1.
Kahai, S. K., Kaserman, D. L., Mayo, J. W., 1996. Is the "dominant firm" dominant? An empirical analysis of AT&T's market power. The Journal of Law and Economics 39 (2), 499–517.
Kam, C. D., Franzese Jr., R. J., 2007. Modeling and Interpreting Interactive Hypotheses in Regression Analysis. University of Michigan Press, Ann Arbor, MI.
Kane, E. J., 2000. Incentives for banking megamergers: What motives might regulators infer from event-study evidence? Journal of Money, Credit and Banking 32 (3), 671–701.
Klein, B., 2001. The Microsoft case: What can a dominant firm do to defend its market position? The Journal of Economic Perspectives 15 (2), 45–62.
Kwoka, J. E., 1985. The Herfindahl index in theory and practice. Antitrust Bulletin 30 (4), 915–947.
Kwoka, J. E., 1998. Herfindahl concentration with an import fringe and with supply constraints. Review of Industrial Organization 13 (4), 401–407.
Lerner, A. P., 1934. The concept of monopoly and the measurement of monopoly power. The Review of Economic Studies 1 (3), 157–175.
Lucas, R. E., 1976. Econometric policy evaluation: A critique. Carnegie-Rochester Conference Series on Public Policy 1 (1), 19–46.
Martin, S., 1988. Market power and/or efficiency? The Review of Economics and Statistics 70 (2), 331–335.

Molyneux, P., Forbes, W., 1995. Market structure and performance in European banking. Applied Economics 27 (2), 155–159.
Radecki, L. J., 1998. The expanding geographic reach of retail banking markets. Economic Policy Review 4 (2), 15–34.
Riordan, M. H., 1998. Anticompetitive vertical integration by a dominant firm. The American Economic Review 88 (5), 1232–1248.
Salop, S. C., Scheffman, D. T., 1983. Raising rivals' costs. The American Economic Review 73 (2), 267–271.
Scherer, F. M., Ross, D., 1990. Industrial Market Structure and Economic Performance, 3rd Edition. Houghton Mifflin Company, Boston, MA.
Shepherd, W. G., 1972. The elements of market structure. The Review of Economics and Statistics 54 (1), 25–37.
Stigler, G. J., 1964. A theory of oligopoly. The Journal of Political Economy 72 (1), 44–61.
Stiroh, K. J., 2000. Compositional dynamics and the performance of the U.S. banking industry. Staff Report 98, Federal Reserve Bank of New York.
Stiroh, K. J., Strahan, P. E., 2003. Competitive dynamics of deregulation: Evidence from U.S. banking. Journal of Money, Credit and Banking 35 (5), 801–828.
Teles, P., Zhou, R., 2005. A stable money demand: Looking for the right monetary aggregate. Economic Perspectives Quarter 1, 50–63.
Tirole, J., 1988. The Theory of Industrial Organization. MIT Press, Cambridge, MA.
White, L. J., 2008. The growing influence of economics and economists on antitrust: An extended discussion. Mimeo 08-3, New York University.

Appendix

Table A.1: Robustness results for the subperiods 1984-1988, 1989-1993, 1994-1998, and 1999-2004

State	θ^*	%*	$year^*$	HHI^*	mfx^*	rents	dollar bonus	n^*	period
Alabama	1.34***	3.13	1984	571.79	2.87	-70.08	-166163.9	7.49	1984-1988
Alabama	0.73***	6.8	1989	918.37	0.32	-50.34	-101180.6	14.8	1989-1993
Alabama	0.25***	9.39	1999	1890.19	0.48	2.87	-22525.63	14.49	1999-2004
Alaska	27.66***	19.44	1994	2911.01	3.17	9.22	3937.05	1.33	1994-1998
Arizona	0.41***	41.6	1999	4598.03	15.41	16.84	14575.49	17.38	1999-2004
Arkansas	0.57***	18.54	1994	135.43	11.39	-6.45	-3051.67	42.46	1994-1998
California	0***	100	1989	1074.26				458.1	1989-1993
California	49.28***	0.17	1994	2676.25	0.68	177.82	29400000	0.59	1994-1998
California	0.12***	36.64	2004	585.56	4.62	11.04	14989.78	77.73	1999-2004
Colorado	0.46***	10.17	1984	234.77	3.52	57.59	11643.14	43.42	1984-1988
Colorado	0.25***	24.95	1989	276.37	88.09	15.06	1521.58	93.89	1989-1993
Colorado	0***	100	1994	318.5				224.53	1994-1998
Connecticut	6.52***	8.22	1984	1828.37	0.64	-27.16	-154322.1	4.8	1984-1988
Connecticut	14.53***	3.03	2001	722.22	19.3	347.37	1368991	0.5	1999-2004
Delaware	6.24***	40.48	1984	1564.34	14.25	31.59	52180.86	6.77	1984-1988
Dist of Columbia	3.31***	27.45	1984	2403.24	0.28	571.73	1304417	5.63	1984-1988
Dist of Columbia	0***	100	1989	2140.34				21.5	1989-1993
Florida	0***	100	1984	364.69				402	1984-1988
Florida	0.02***	86.95	1989	372.59	0.5			345.05	1989-1993
Florida	0.06***	88.7	1999	295.62	0.65			231.43	1999-2004
Georgia	0.13***	48.91	1997	916.31	0.43	1.97	-1201.19	113.99	1994-1998
Georgia	3.46***	0.41	1999	1219.29	12.53	41.45	1856726	1.34	1999-2004
Hawaii	0***	100	1989	3166.74				19.38	1989-1993
Hawaii	25.72***	26.67	1999	4038.03	0.41	-169.2	-1266790	2	1999-2004
Idaho	4.4***	16.67	1984	2326.83	114.61	96.38	142775.9	4	1984-1988
Idaho	52.37***	0						0	1989-1993
Illinois	0***	100	1984	678.95				1204.26	1984-1988
Illinois	0***	100	1989	674.91				1036.87	1989-1993
Indiana	0.23***	26.86	1984	209.35	0.9	-2.51	-2285.19	95.06	1984-1988
Indiana	0***	100	1989	196.98				279.54	1989-1993
Indiana	0.43***	20.91	1994	504.18	1.13	12.97	7316.91	40.9	1994-1998
Indiana	0***	100	1999	921.53				150.72	1999-2004
Iowa	0.05***	88.37	1994	53.82	3.64			414.85	1994-1998
Iowa	0.02***	99.23	1999	279.35	9.25			408.83	1999-2004
Kansas	0.19***	26.83	1991	259.36	0.84	-0.11	-505.09	138.39	1989-1993
Kansas	0.19***	31.62	1994	381.22	1.86	-1.48	-847.35	131.71	1994-1998
Kansas	0.5***	10.08	1999	242.84	3.86	6.43	320.4	36.9	1999-2004
Kentucky	0.54***	14.15	2003	462.81	3.93	35.17	25937.82	25.14	1999-2004
Louisiana	0.53***	10.92	1984	175.18		-64.98	-41847.2	30.61	1984-1988
Louisiana	0***	100	1989	669.93				222.55	1989-1993
Louisiana	0.65***	9.75	1994	640.56	3.64	-2.74	-12214.73	16.96	1994-1998
Maine	19.66***	6.98	1985	1177.67	0.03	-0.45	-10027.47	1.09	1984-1988
Maine	87.09***	0						0	1999-2004
Massachusetts	0***	100	1984	1230.37				110.78	1984-1988
Michigan	0***	100	1985	629.95				325.59	1984-1988
Michigan	0***	100	1999	1285.63				162.22	1999-2004
Minnesota	0***	100	1999	2029.77				476.11	1999-2004
Mississippi	0.3***	48.36	1989	748.07	258.61	-26.13	-9528.83	58.79	1989-1993
Missouri	3.25***	0.92	1986	236.08	0.97	4.89	-1164.73	4.28	1984-1988
Missouri	0.15***	19.85	1994	423.85	1.09	1.16	-1880.94	84.34	1994-1998

Continued on next page ...

Table A.1 (Continued from previous page)

State	θ^*	%*	*year**	*HHI**	*mfx**	rents	dollar bonus	n^*	period
Missouri	0.18***	26.96	2001	434.45	0.34	10.7	2398.52	89.53	1999-2004
Montana	0.44***	52.25	1991	171.42	7.51	12.92	655.94	63.07	1989-1993
Montana	2.8***	4.63	1994	301.88	6.09	-13.53	-12786.46	4.6	1994-1998
Montana	22.14***	1.27	2001	847.07	0.31	57.89	127116.5	0.65	1999-2004
Nebraska	3.28***	0.7	1989	252.53	8.79	-25.58	-38972.15	2.6	1989-1993
Nebraska	0.81***	7.65	1999	207.77	4.53	7.1	708.95	20.15	1999-2004
Nevada	2.19***	23.75	1999	2600.93	1.41	2.22	-1841.65	6.51	1999-2004
New Hampshire	0***	100	1984	391.69				54.84	1984-1988
New Hampshire	4.04***	38.24	2000	4312.38	8.15	20.4	11615.03	4.65	1999-2004
New Jersey	0***	100	1994	766.48				76.26	1994-1998
New Jersey	0***	100	1999	1587.46				78.95	1999-2004
New Mexico	0.49***	59.52	1985	466.91	1.49	43.91	5351.25	56.01	1984-1988
New Mexico	2.34***	7.94	1989	538.51	0.69	-67.74	-47523.98	6.79	1989-1993
New Mexico	0.52***	60	1997	879.3	7.17	28.28	5799.18	39.3	1994-1998
New York	0***	100	1984	1082.75				192.82	1984-1988
North Carolina	1***	11.78	1989	1717.66	35.46	35.39	238042.9	9	1989-1993
North Dakota	1.02***	12.22	1984	122.08	40.9	-0.81	-763.02	21.12	1984-1988
North Dakota	0.16***	86.42	1999	394.81	1.32	5.36	46.02	90.42	1999-2004
Ohio	0.06***	47.25	1994	701.1	0.18			114.54	1994-1998
Oklahoma	0.35***	13.01	1992	181.69	1.33	18.11	3395.75	50.25	1989-1993
Oklahoma	8.92***	0.52	1996	374.18	0.67	16.17	43036.85	1.39	1994-1998
Oregon	0***	100	1989	3062.33				45.99	1989-1993
Oregon	0.87***	43.18	1999	716.36	5.39	22.94	12667.51	16.22	1999-2004
Pennsylvania	1.8***	3.29	1989	472.97	47.57	43.36	378414.7	9.38	1989-1993
Pennsylvania	0.06***	61.1	1994	1115.12	0.24			134.19	1994-1998
Rhode Island	12.58***	15	1984	4660.74	197	32.89	147552.2	1.78	1984-1988
Rhode Island	0***	100	1989	4098.08				10.93	1989-1993
Rhode Island	0.16***	91.43	1994	5857.59	35.07	165.84	518790.1	6.49	1994-1998
South Dakota	2.15***	3.33	1987	838.17	0.75	23.33	14299.1	4	1984-1988
South Dakota	5.86***	0.88	1989	1045.58	6.41	117.4	248054	1	1989-1993
South Dakota	0.4***	41.18	1994	1700.23	0.89	-7.88	-2392.78	42.26	1994-1998
South Dakota	0***	100	1999	2326.59				83.28	1999-2004
Tennessee	11.22***	0.35	1984	400.97	0.04	37.26	167580.9	1	1984-1988
Tennessee	0.17***	27.97	1999	1766.69	0.28	5.33	-466.05	53.88	1999-2004
Texas	0.03***	40.71	1994	692.22	0.48			360.06	1994-1998
Texas	0.16***	15.2	2001	312.37	32.59	12.92	5499.5	89.5	1999-2004
Utah	26.16***	2.24	1985	1659.08	0.43	83.34	224403.5	1.2	1984-1988
Utah	5.44***	7.43	1994	2009.21	26.05	-15.8	-94304.52	2.96	1994-1998
Utah	39.44***	1.95	1999	2513.3	11.13	41.2	883668.8	1	1999-2004
Vermont	8.06***	19.38	1984	1038.7	0.27	17.44	7657.46	5	1984-1988
Vermont	1.01***	86.54	1994	1149.8	0.78	5.12	-158.48	18.02	1994-1998
Virginia	0.32***	16.77	1984	1162.51	0.7	0.33	-5245.88	28.38	1984-1988
Virginia	0.45***	11.43	1989	1292.42	0.6	-10.8	-31761.7	19.36	1989-1993
Washington	7.05***	4.09	1984	1710.37	30.11	35.2	101090.6	3.81	1984-1988
Washington	15.14***	2.22	1994	1514.19	2.98	19.76	101803.8	1.84	1994-1998
Washington	0.58***	44.8	1999	433.39	0.43	-4.2	-4059.43	35.17	1999-2004
West Virginia	0.44***	31.14	1984	101.41	0.58	-16.45	-2939.14	66.21	1984-1988

Results are based on the preferred specification among equations (9a)(9c) as discussed in the text. Also, results are reported for only states and sub-periods in which evidence of collusion is obtained. Significance is indicated at the following levels: 1/5/10% (***/**/*, respectively).

Table A.2: Omitted variable bias

$[A]$ $L_{i,t}, \theta_{i,t}$	$[B]$ $L_{i,t}$, HHI_t	$[C\ ]$ L_t, HHI_t (Cowling and Waterson, 1976)
$$L_{i,t} = \beta_i + \beta_1 \theta_{i,t} + \epsilon_{i,t} \quad \text{(A.1)}$$ • Include market share ($\theta_{i,t}$) as a proxy for market power; • Firm-level markup.	$$L_{i,t} = \beta_i + \beta_1 HHI_t + \epsilon_{i,t} \quad \text{(B.1)}$$ • Include the Hirschman-Herfindahl Index (HHI) as a proxy for market power; • Firm-level markup.	$$L_t = \beta_0 + \beta_1 HHI_t + \epsilon_t \quad \text{(C.1)}$$ • Include the Hirschman-Herfindahl Index (HHI) as a proxy for market power; • Industry-level markup.
$$\lambda_{i,t} = \gamma_i + \gamma_1 \theta_{i,t} + w_{i,t} \quad \text{(A.2)}$$ • $\lambda_{i,t}$ is a function of $\theta_{i,t}$ in an auxiliary regression; • $\lambda_{i,t}$ is also a function of a firm-specific effect γ_i.	$$\lambda_{i,t} = \frac{\gamma}{\theta_{i,t}} + \frac{\gamma_1}{\theta_{i,t}} HHI_t - \frac{\sum_{j=1}^{N} \theta_{j,t} \lambda_{j,t}}{\theta_{i,t}} + w_{i,t}, j \neq i \quad \text{(B.2)}$$ • $\lambda_{i,t}$ is a function of $\theta_{i,t}$ in an auxiliary regression: $\lambda_{i,t} = \gamma_i + \gamma_1 \theta_{i,t} + \omega_{i,t}$; • $\lambda_{i,t}$ is also a function of a firm-specific effect γ_i • Multiplying both sides with $\theta_{i,t}$, and aggregating across all firms, we arrive at: $\theta_{i,t}\lambda_{i,t} + \sum_{j=1}^{N} \theta_{j,t}\lambda_{j,t} = \gamma + \gamma_1 HHI_t + w_{i,t}, j \neq i$ • Rearranging to solve for $\lambda_{i,t}$ gives equation (B.2).	$$\lambda_t = \gamma + \gamma_1 HHI_t + \omega_t \quad \text{(C.2)}$$ • λ_t is a function of HHI_t in an auxiliary regression; • λ_t can be a function of an industry-specific effect (not here).
$$L_{i,t} = [\beta_i + \beta_2\gamma_i] + \left[(\beta_1 + \gamma_1\beta_2)\theta_{i,t}\right] + \left[\beta_2 w_{i,t} + \epsilon_{i,t}\right] \quad \text{(A.3)}$$ • Equation (A.1) after introducing $\lambda_{i,t}$ using equation (A.2).	$$L_{i,t} = \left[\beta_i + \beta_2 \frac{\gamma}{\theta_{i,t}}\right] + \left[\left(\beta_1 + \beta_2 \frac{\gamma_1}{\theta_{i,t}}\right) HHI_t\right] - \beta_2 \frac{\sum_{j=1}^{N} \theta_{j,t}\lambda_{j,t}}{\theta_{i,t}} + \left[\beta_2 w_{i,t} + \epsilon_{i,t}\right] \quad \text{(B.3)}$$ • Equation (B.1) after introducing $\lambda_{i,t}$ using equation (B.2).	$$L_t = [\beta_0 + \beta_2\gamma] + [(\beta_1 + \beta_2\gamma_1) HHI_t] + [\beta_2 w_t + \epsilon_t] \quad \text{(C.3)}$$ • Equation (C.1) after introducing $\lambda_{i,t}$ using equation (C.2).
$$E(\hat{\beta_1}) = \beta_1 + \beta_2\gamma_1 = \beta_1 + \beta_2 \left[\frac{Cov\left[\theta_{i,t}, \lambda_{i,t}\right]}{Var\left[\theta_{i,t}\right]}\right] \quad \text{(A.4)}$$ • Coefficient for $\theta_{i,t}$; • $Cov\left[\theta_{i,t}, \lambda_{i,t}\right] > 0$ from Stigler (1964) and Section 2.2; • The higher $Var\left[\theta\right]$, the less likely collusion is; • β_2 is (expected to be) positive and increasing in the amount of collusion, based on Cowling (1976) and Cowling and Waterson (1976).	$$E(\hat{\beta_1}) = \beta_1 + \beta_2 \left[\frac{\gamma_1}{\theta_{i,t}}\right] = \beta_1 + \beta_2 \left[\frac{\frac{Cov[\theta_{i,t}, \lambda_{i,t}]}{Var[\theta_{i,t}]}}{\theta_{i,t}}\right] \quad \text{(B.4)}$$ • Coefficient for HHI_t; • $Cov\left[\theta_{i,t}, \lambda_{i,t}\right] > 0$; • For a given $Cov\left[\theta_{i,t}, \lambda_{i,t}\right] > 0$, as $\theta_{i,t}$ increases by h, $Var[\theta_{i,t}]$ increases by h^2; • So the size of β_2 is uncertain, but likely to drop quickly in highly concentrated markets.	$$E(\hat{\beta_1}) = \beta_1 + \beta_2\gamma_1 = \beta_1 + \beta_2 \left[\frac{Cov[HHI_t, \lambda_t]}{Var[HHI_t]}\right] \quad \text{(C.4)}$$ • Coefficient for HHI_t; • $Cov\left[HHI_t, \lambda_t\right] > 0$ from Stigler (1964) and Section 2.2; • The $Var\left[HHI_t\right]$ is expected to be relatively small, except (perhaps) across industries; • β_2 is (expected to be) positive, based on Cowling (1976) and Cowling and Waterson (1976).